VALUE IN HEALTH 18 (2015) 467–476

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Health Policy Analysis

Assessment of the Quality of the Clinical Evidence in Submissions to the Australian Pharmaceutical Benefits Advisory Committee: Fit for Purpose?

Michael Wonder, BSc (Hons), BPharm[1,*], *Sheryl Dunlop, MSc*[2]

[1]*Wonder Drug Consulting Pty. Ltd., Cronulla, NSW, Australia;* [2]*Market Access, Asia-Pacific, Zimmer Pty Ltd, Belrose, NSW, Australia*

ABSTRACT

Background: Assessments of the comparative clinical (and cost) effectiveness of new medicines are increasingly being used to inform decisions on their reimbursement. Assessments of added clinical benefit are invariably based on evidence generated to support registration. **Objective:** Our objective was to identify and characterize significant problems relating to the quality of the clinical evidence in submissions to the Australian Pharmaceutical Benefits Advisory Committee (PBAC) seeking subsidy on the Pharmaceutical Benefits Scheme and thus determine whether the evidence presented to the committee was "fit for purpose." **Methods:** We conducted a retrospective analysis of submissions considered by the PBAC between 2005 and 2012 using a published evaluation framework. We developed an additional framework to categorize significant problems in more detail. Significant problems related to the choice of comparator, the unavailability of randomized clinical trial evidence, poor-quality data, a claim of clinical superiority, and a claim of clinical noninferiority. **Results:** We identified 261 significant problems in 479 major submissions. There was a significant problem with the sponsor's choice of comparator in 11% of the submissions. The most common significant problem (29%) was the determination of a medicine's comparative performance in the target patient population. **Conclusions:** The supporting clinical evidence is the foundation of a PBAC submission. We found a poor fit for purpose; on average, one in every two major submissions had a significant problem with the supporting evidence. The findings from our study, if confirmed in other jurisdictions, raise important questions regarding what clinical evidence should be generated to support the reimbursement of new medicines.

Keywords: decision making, evidence, quality, reimbursement.

Introduction

Governments of the developed world currently face challenges in ensuring that their constituents are able to access new and effective health care technologies in a timely and affordable manner. They have promoted the use of health technology assessment (HTA) to facilitate efficient use of their limited public resources. Although some undertake assessments of comparative economic effectiveness ("value for money"), a common denominator for all is an assessment of comparative clinical effectiveness ("level of added clinical benefit") [1]. Assessments of added clinical benefit, be they single or multiple and direct or indirect, are invariably based on evidence generated to support registration. Few have studied whether the clinical evidence generated to support the registration of new medicines is well suited for reimbursement/coverage decision making. Insofar as access to new medicines is becoming increasingly dependent on reimbursement, this is an important public health issue.

Australia has considerable experience in the use of HTA to inform reimbursement decision making. The Pharmaceutical Benefits Scheme (PBS) was established in 1953 under the National Health Act to guarantee Australians subsidized access to essential medicines. The National Health Act also established the Pharmaceutical Benefits Advisory Committee (PBAC) to make recommendations to the Commonwealth Minister for Health regarding the subsidy of medicines on the PBS. The PBAC has 20 years experience in assessing submissions to list new medicines on the PBS or make a substantial change to currently listed medicines (so-called major submissions) in terms of their comparative clinical and economic effectiveness [2].

The main objective of our study was to identity and then characterize significant problems relating to the quality of the clinical evidence in major submissions to the PBAC using the evaluation framework of Hill et al. [3] with information on the submissions from their public summary documents (PSDs) [4].

Conflicts of interest: The views expressed in this article do not necessarily reflect the views or practices of our previous and/or current employers.

* *Address correspondence to*: Michael Wonder, PO Box 470, Cronulla, NSW, Australia 2230.

E-mail: wonderdrug@optusnet.com.au.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2015.02.011

We wanted to determine the type of problems that had a bearing on PBAC's decision making and thus determine whether the clinical evidence presented to the committee was "fit for purpose."

Methods

Because the focus of our study was the quality of the clinical evidence, aspects not directly related to this (i.e., modeling issues, calculation errors, and administrative matters) were not considered.

We deemed a problem to be significant if the issue was serious enough to prevent the PBAC from making a recommendation for the medicine in accordance with the request in the submission at the time. The reason(s) why the PBAC made a decision to recommend (or not recommend) the listing of a medicine is documented in the associated PSD. This does not suggest that there were no significant problems with submissions for medicines that were recommended.

We felt that it was important to distinguish between "significant problems" and "uncertainty." Submissions will always be associated with uncertainty, even those that are recommended. We sought to identify situations in which the level of uncertainty in submissions was so great that it presented a significant problem to the PBAC.

The secrecy provisions of the National Health Act bind the PBAC and submissions are treated as "commercial in confidence." The signing of the Australia-United States Free Trade Agreement in early 2005 facilitated the release of further information regarding the basis for PBAC's determinations regarding the subsidy of medicines on the PBS in PSDs from mid-2005 [2].

PSDs are available only for submissions related to PBAC considerations on the listing of medicines; they are not available for other submissions for PBAC considerations, such as those relating to pricing arrangements for listed medicines. We included all published PSDs associated with submissions (initial submissions and resubmissions) for medicines and vaccines seeking a listing on the PBS. We excluded submissions with PSDs for the following:

- Vaccines seeking a listing on the National Immunisation Program [5].
- Fixed-dose combination products seeking a listing on the PBS.
- Medicines seeking a listing on the Life Saving Drugs Program [6].
- Medicinal preparations (e.g., nutritional supplements) or devices seeking a listing on the PBS.
- New strengths or formulations of medicines already listed on the PBS.
- Nonprescription medicines seeking a listing on the PBS.
- Medicines seeking a change to an existing therapeutic relativity to another listed medicine.
- Medicines for which the applicant was not the medicine's sponsor.

The largest proportions of the submissions we excluded were for fixed-dose combination products, National Immunisation Program vaccines, and Life Saving Drugs Program medicines. Insofar as the (current) PBAC guidelines consider fixed-dose combination products and National Immunisation Program vaccines as separate product types, they are subject to different evidence requirements and hence their exclusion from the analysis is justified. The exclusion of submissions for Life Saving Drugs Program medicines is justified insofar as they are also subject to additional decision-making criteria.[6]

We sought to identify the types of problems in submissions as used by Hill et al. [3] that had been first described by O'Brien [7] and are summarized in Table 1.

Hill et al. found that problems with the supporting clinical evidence were the most common of all problem categories, but they did not conduct further analysis to obtain deeper insights. We developed additional frameworks in an attempt to understand these problems at their core (Tables 2 and 3).

One of us (M.J.W.) developed a coding template. We coded each eligible PSD using the template independently; differences in opinion were resolved by consensus.

Some PSDs were for submissions with multiple requests that were associated with different target patient populations (e.g., treatment-naive and treatment- experienced patients), different proposed main comparators, and different clinical claims. In these situations, we examined each request for each patient population because there might have been a significant problem with one request but not the other.

Some submissions were associated with requests in the form of options. We examined all options because there may have been significant problems in those that were not accepted by the PBAC.

Determinations for all problem categories were made on the basis of the clinical evidence presented in the submission. In situations in which the submission did not include any clinical data for PBAC's preferred main comparator, a determination on the estimate of comparative clinical efficacy could not be made.

Results

The PBAC published 598 PSDs for submissions considered between July 2005 and November 2012; we excluded 119 (20%) submissions, resulting in a study sample of 479 (80%) submissions (Fig. 1).

Some submissions were excluded for more than one reason. We identified 261 significant problems in the 479 submissions, an average of 0.54 significant problems per submission (Table 4). The 479 submissions were associated with 483 PBAC outcomes. Eighty-two percent of the submissions with a significant problem were associated with a rejection by the PBAC. Some submissions were recommended despite having one or more problems with the supporting clinical evidence. Invariably, they were recommended on a different clinical and economic basis than proposed by the sponsor. Submissions with no major problems with the supporting clinical data might have been rejected by the PBAC for another reason, such as uncertain or unacceptable cost-effectiveness.

There was a significant problem with the sponsor's choice of comparator in 11% of the submissions. There was no clear temporal pattern, with at least one problem occurring at all bar two PBAC meetings.

Randomized clinical trial evidence was not available for 4% of the submissions. The most common significant problem (140 or 29% of all submissions) was the determination of the medicine's comparative performance in the target patient population. Some submissions had multiple problems insofar as they contained multiple comparisons with different clinical claims (i.e., a claim of clinical superiority vs. comparator A and a claim of clinical noninferiority vs. comparator B).

There were a few examples in which the initial submission did not clearly identify the target patient population for the proposed medicine, and it took one to two resubmissions to resolve this problem.

Most of the problems related to the medicine's comparative performance with respect to efficacy, with only a few examples

Table 1 – Categorization of the significant problems with the clinical evidence in major submissions to the PBAC.

Problem area	Specific problem	PBAC's position	Detail of problem	Coding issues and challenges
Choice of comparator	Choice of main comparator	*The main comparator* is defined as the therapy that prescribers would most replace with the proposed drug in practice if the PBS subsidizes the proposed drug as requested. In practice, however, the main comparator can be difficult to identify.	The PBAC was of the view that the proposed main comparator in the submission was inappropriate or that additional comparisons were required.	A submission was deemed to have a significant problem if the PBAC had a concern about the sponsor's choice of the main or secondary comparator(s) to a sufficient level that the submission was rejected on that basis. If the submission did not include any clinical evidence on the PBAC's preferred comparator, then it was not possible for us to make assessments on other aspects in the framework
Estimate of comparative clinical efficacy	Availability of randomized trial evidence	The PBAC has a strong preference for clinical and economic evaluations that are based on direct randomized trials; i.e., trials that directly compare the proposed drug with the main comparator.	The PBAC has considered and will continue to consider all levels of evidence. However, the PBAC will be most influenced by the results of direct randomized trials as the most rigorous source of data.	The supporting clinical evidence in submissions ranged from case reports/case series to meta-analyses of multiple randomized trials. For some trials, we needed to use PubMed to determine whether treatment allocation was conducted in a randomized manner. A submission was classified as having a significant problem if there was no supporting randomized clinical trial evidence, irrespective of whether or not the trial was a direct one
	Poor-quality evidence	The purpose of assessments of measures to minimize bias is to provide the sponsor and the PBAC with a clear idea of which trials are of greater scientific rigor. There is no minimum standard, but the PBAC is most likely to be persuaded by the data of the highest scientific rigor.	The PBAC has developed a quality checklist for the assessment of the quality of the randomized trial evidence. The checklist covers randomization, blinding, and follow-up.	There are many aspects to poor-quality clinical evidence. The definition of poor-quality clinical evidence used by Hill et al. [3] is not clearly stated in their article. To avoid double counting, we did not consider problems associated with treatment allocation. Insofar as poor-quality clinical evidence presents a major challenge in the determination of an estimate of comparative clinical efficacy, we were mindful of further double counting. We considered a submission to have a significant problem if the PBAC made an explicit statement in section 12 of the submission's PSD on the poor quality of the supporting clinical evidence. We recognize that our count for problems in this category may be lower than otherwise expected
	Use of surrogate outcome	The direct randomized trials might report only those outcomes that are of less patient relevance than intended final outcomes of treatment. These less relevant outcomes are known as *surrogate outcomes*. Arguably, the closer a surrogate outcome is to the final outcome, the more useful it is, but	To transform the surrogate outcomes measured in the trials to final outcomes and to extend the range of outcomes, the trial results might need to be supplemented by estimates obtained from other sources	Insofar as the use of a problematic surrogate outcome in a submission would have presented a challenge to the PBAC regarding the determination of the magnitude of the clinical benefit of the proposed treatment in the target patient population on final outcome, there is a risk of double counting. Because the PBAC has not published a list of acceptable surrogate outcomes for use in PBAC

continued on next page

Table 1 – *continued*

Problem area	Specific problem	PBAC's position	Detail of problem	Coding issues and challenges
		generally the more difficult it is to measure accurately.		submissions, we chose not to code the submissions on this category.
	Analysis of interpretation of clinical evidence	The interpretation of the clinical data presented in a submission is crucial in determining its success. The PBAC has developed a framework for the classification of the therapeutic relativity of the proposed medicine over its main comparator.	The essential difference between assessing whether the proposed medicine is superior to the main comparator is that the 95% confidence interval for superiority excludes the possibility that there is no difference between the two treatments.	This category relates to submissions for which a claim of clinical superiority was made and goes to the quality, strength, and relevance of the supporting clinical evidence. There is some overlap with poor-quality evidence. Although many submissions had clinical problems, some of them did not relate to the supporting clinical evidence (i.e., the PBAC was uncertain about the clinical need for the medicine or the PBAC had some concerns about the feasibility of the proposed restrictions). These nonclinical evidence issues were not considered to be "significant problems"
	Determination of therapeutic noninferiority	The interpretation of the clinical data presented in a submission is crucial in determining its success. The PBAC has developed a framework for the classification of the therapeutic relativity of the proposed medicine over its main comparator.	The essential difference between assessing whether the proposed medicine is noninferior to the main comparator is that the 95% confidence interval for noninferiority excludes the possibility that the proposed medicine is inferior to a clinically important extent.	This category relates to submissions for which a clinical claim of at least noninferiority was made. Here, a submission was deemed to have a significant problem if • The PBAC was satisfied that the proposed medicine is noninferior to the comparator(s) but at a different therapeutic relativity to that proposed in the submission and the submission was recommended on this basis (data interpretation issue). • The PBAC was not satisfied that the proposed medicine is noninferior to the proposed comparator(s) because of a wide noninferiority margin and the submission was rejected on that basis (data quality issue). • The PBAC was not satisfied that the proposed medicine is noninferior to the comparator(s) and the submission was rejected or deferred on this basis (data interpretation issue). • The PBAC was satisfied that the proposed medicine is actually inferior to the comparator (s) and the submission was recommended on that basis (i.e., less effective but also less costly) (data interpretation issue). • The PBAC wanted to obtain further clinical advice.

PBAC, Pharmaceutical Benefits Advisory Committee; PBS, Pharmaceutical Benefits Scheme; PSD, public summary document.

Table 2 – Categorization of the significant problems with submissions to the PBAC that included a claim of clinical superiority.

Scenario	Proposed comparator	Target patient population	Proposed outcome	Treatment effect size	Clinical significance	Clinical evidence issue	Justification	PBAC outcome	Problem category
A	Accepted	Unclear (no evidence)	NA	NA	NA	The clinical evidence in the submission does not encompass the target patient population	The PBAC made the comment in the PSD that the target patient population was not identified in the submission and a resubmission was subsequently prepared and lodged with different clinical evidence base or a resubmission could not be prepared and lodged as the requisite data do not exist.	Rejection	Data relevance
B	Accepted	Unclear (evidence not identified)	NA	NA	NA	The target patient population is an unidentified and unanalyzed subgroup of the patient population of the clinical evidence	The PBAC made the comment in the PSD that the target patient population had not been identified in the submission. A resubmission was subsequently prepared and lodged with the same clinical evidence base but with the target patient population being a clearly specified subgroup of the patient population of the initial submission.	Rejection	Data relevance
C	Accepted	Clear	Not accepted	NA	NA	Relationship of the proposed outcome to a patient-relevant final outcome not established in a quantitative and/or quantitative manner	The PBAC made the comment in the PSD that the proposed clinical outcome is unacceptable or the proposed outcome is acceptable but the submission contained insufficient evidence to support the proposition that the outcome is related to a patient-	Rejection	Data quality

continued on next page

Table 2 – *continued*

Scenario	Proposed comparator	Target patient population	Proposed outcome	Treatment effect size	Clinical significance	Clinical evidence issue	Justification	PBAC outcome	Problem category
D	Accepted	Clear	Accepted	Unclear	NA	Trial design and conduct shortcomings	relevant final outcome or another patient-relevant outcome. The PBAC made the comment in the PSD that the proposed outcome is acceptable and that its relationship to a patient-relevant final outcome is sufficiently well established but the magnitude of the treatment effect on a patient- relevant final outcome is unclear.	Rejection	Data quality
E	Accepted	Clear	Accepted	Accepted	Unclear/not established	Minimal clinically important difference not established	The PBAC made the comment in the PSD that the proposed outcome is acceptable and that the magnitude of the treatment effect on a patient-relevant outcome is clear but the clinical significance of the treatment effect size is uncertain.	Recommendation, rejection, or deferral	Data interpretation
F	Accepted	Clear	Accepted	Unclear	Unclear	The PBAC disputed the claimed clinical benefit	The PBAC made the comment in the PSD that the clinical claim is unreasonable and that a different clinical claim is appropriate.	Recommendation, rejection, or deferral	Data interpretation

NA, not applicable; PBAC, Pharmaceutical Benefits Advisory Committee; PSD, public summary document.

Table 3 – Categorization of the significant problems with submissions to the PBAC that included a claim of clinical noninferiority.

Scenario	Claim of therapeutic noninferiority	Clinical evidence issue	Justification	PBAC outcome	Problem category
M	Not accepted	The PBAC was unable to determine whether the medicine is noninferior or inferior to comparator with respect to efficacy and/or safety.	The PBAC made the comment in the PSD that the clinical claim in the submission was unreasonable and that it was unable to determine what clinical claim could be made.	Rejection	Data quality
N	Not accepted	The PBAC was satisfied that the clinical evidence in the submission supports a claim of inferiority (efficacy) or a different position (safety)	The PBAC made the comment in the PSD that a different clinical claim was reasonable.	Recommendation (different basis)	Data interpretation
O	Not accepted	Noninferiority (efficacy) margin has not been established or was not specified in the clinical trial.	The PBAC made the comment in the PSD that the noninferiority margin was unacceptably wide.	Rejection	Data quality
P	Not accepted	The PBAC was satisfied that the clinical evidence in the submission supports a different therapeutic relativity to that proposed by the sponsor.	The PBAC made the comment in the PSD that the proposed therapeutic relativity was unreasonable and that a different therapeutic relativity was appropriate.	Recommendation (different therapeutic relativity)	Data interpretation
Q	Not accepted	The PBAC was not clear about the clinical positioning of the medicine or was uncertain that the clinical evidence supported the clinical claim.	The PBAC made the comment in the PSD that it wanted to obtain further clinical advice.	Deferral or rejection	Data interpretation

PBAC, Pharmaceutical Benefits Advisory Committee; PSD, public summary document.

related to a comparative safety claim (e.g., a claim of a lower incidence of hypoglycemic events).

A common scenario was that the PBAC did not accept the claim of clinical superiority over the comparator because of uncertainty over the magnitude of the incremental treatment effect on a final outcome.

Matters of interpretation such as the determination of the clinical relevance of the difference of the proposed medicine over its comparator on an acceptable outcome were often resolved with further discussion and resubmission(s) without the need for substantive new clinical evidence. These submissions often had pricing implications.

The determination of clinical noninferiority was a problem in 34 (7%) submissions; all related to a claim of noninferiority that was inadequately supported or justified. In some cases, the submission was rejected, whereas in others the submission was recommended on the basis of a different therapeutic relativity. In three cases, the medicine was recommended on the basis of a different clinical claim (i.e., clinical inferiority) (Table 5).

Discussion

The supporting clinical evidence is the foundation of a robust and credible PBAC submission. We found a rather poor fit for purpose; in recent times, on average, one in every two major submissions to the PBAC had a significant problem with the supporting clinical evidence. A problem with the choice of comparator occurred, on average, with 1 in every 10 major submissions.

The absence of randomized controlled trial (RCT) evidence in a submission was not a frequent occurrence (4% of all submissions) and was seldom a reason for its rejection. In our view, the absence of RCT evidence in a submission is not a significant problem. We share the concern of past National Institute for Health and Care Excellence Chairman Sir Michael Rawlins about RCTs being considered by some to be the “gold standard” of clinical evidence. Although RCTs are very important, evidence from other study designs “can be appropriate in some circumstances” [8].

There are many aspects to poor-quality clinical evidence (e.g., inadequate blinding, suboptimal dose of comparator, early crossover, inadequate wash-out period, high discontinuation rate, and short period of observation). We could find an explicit mention in the PSDs of the supporting clinical evidence being of poor quality in only 15 (3%) submissions; this figure seems artificially low given that there were 68 submissions for orphan drugs. We chose not to consider the use of surrogate outcome as a problem because the use of a surrogate outcome in a submission is in itself not a reason for its rejection and the PBAC is yet to publish a list of acceptable (or unacceptable) surrogate outcomes. Nonetheless, our research shows that the use of a surrogate outcome in a submission did not always pose a significant problem to the PBAC.

We developed an additional evaluation framework to investigate the apparent shortcomings of the supporting clinical

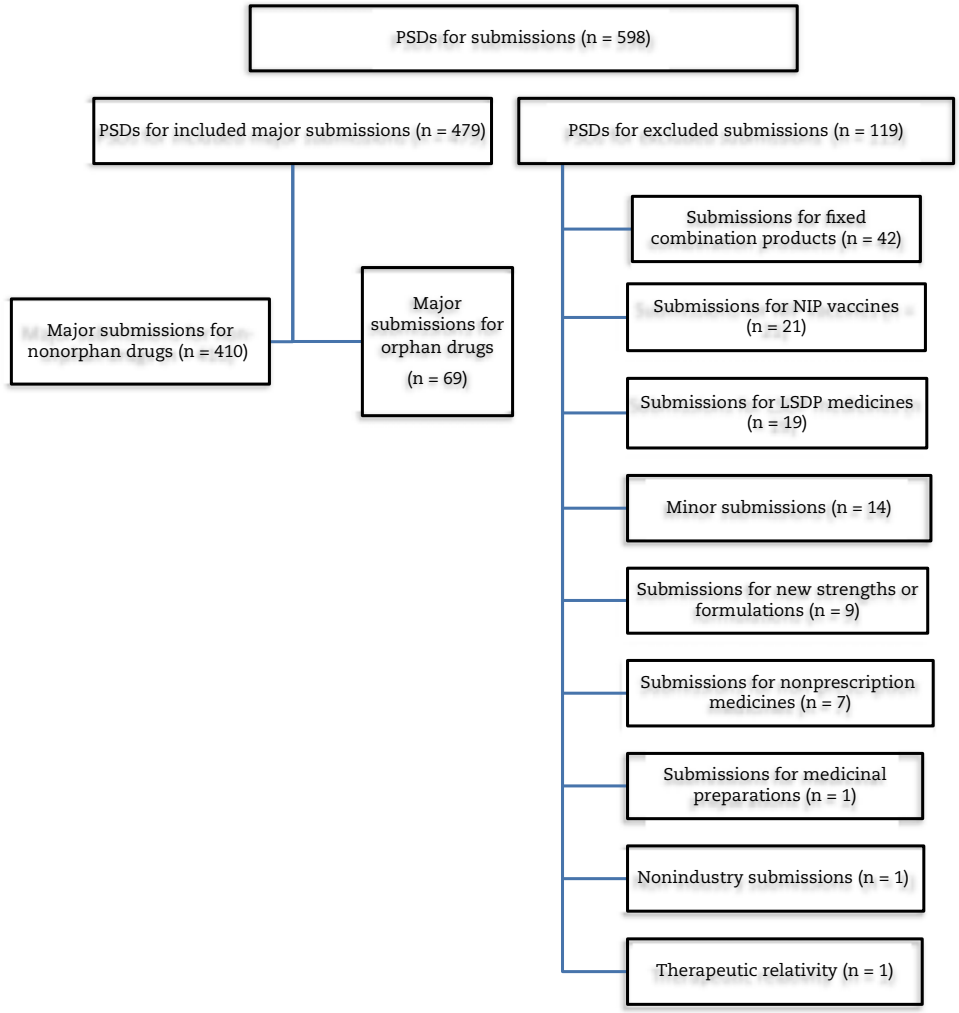

Fig. 1 – Included and excluded submissions. LSDP, Life Saving Drugs Program; NIP, National Immunisation Program; PSD, public summary document.

evidence. We found that most of the supporting clinical evidence was relevant to the reimbursement decision-making question. The results clearly show that the most frequent problem category was the inadequacy of the clinical evidence in being able to support a claim of clinical superiority and/or clinical noninferiority. In the case of the former, it shows that the PBAC often had to grapple with a claim of clinical superiority on a patient-relevant final outcome based on a favorable change in an acceptable surrogate outcome. In the case of the latter, it shows that submissions that included a seemingly safer claim of clinical noninferiority were not without their problems, with the PBAC being unsure as to whether a claim of clinical inferiority might be the correct conclusion.

Our study has a number of limitations. We identified a small number of major submissions with no associated PSD. Their exclusion is unlikely to have a bearing on our results. We excluded 107 (18%) submissions for various reasons (Fig. 1). The content of a given PSD is negotiated between the PBAC and the medicine's sponsor, so it is possible that information on some problems has not been published. Any omitted problems are likely to have been minor ones that did not have a major influence on a submission's outcome.

Table 4 – Categorization of the types of significant problems identified in submissions to the PBAC (July 2005–November 2012).

Problem area	Specific problem	No. (%) of all problems
Choice of comparator	Choice of main comparator	53 (20)
Estimate of comparative clinical efficacy	Availability of randomized clinical trial evidence	18 (7)
	Poor-quality evidence	16 (6)
	Analysis of interpretation of clinical evidence	140 (54)
	Determination of therapeutic noninferiority	34 (13)

PBAC, Pharmaceutical Benefits Advisory Committee.

Table 5 – Submissions with significant problems with the associated claim of clinical superiority or noninferiority.

Submission claim	Problem category	Description	No. (%) of all problems
Clinical superiority	A (Data relevance)	No clinical data for target patient population	5 (4)
	B (Data relevance)	Clinical data for target patient population inadequately identified	10 (7)
	C (Data quality)	Proposed surrogate outcome not accepted	10 (7)
	D (Data quality)	Treatment effect on final outcome not clear	85 (61)
	E (Data interpretation)	Clinical significance of treatment effect on final outcome in dispute	6 (4)
	F (Data interpretation)	The PBAC is of the view that a different clinical claim is more appropriate	24 (17)
Clinical noninferiority	M (Data quality)	The PBAC was unable to determine a therapeutic relativity	26 (76)
	N (Data interpretation)	The PBAC was of the view that a different clinical claim was more appropriate	3 (9)
	O (Data quality)	The PBAC had concerns about the noninferiority margin	1 (3)
	P (Data interpretation)	The PBAC was of the view that a different therapeutic relativity was more appropriate	2 (6)
	Q (Data interpretation)	The PBAC wanted to obtain further clinical advice	2 (6)

PBAC, Pharmaceutical Benefits Advisory Committee.

In a landmark article published in *JAMA* in 2000, Hill et al. [3] reported on their review of 326 major submissions to the PBAC from 1994 to 1997. They identified many significant problems; of a total of 326 submissions, 218 had serious problems of interpretation and were included in their analysis. Significant problems were those considered to have a significant bearing on the decision making of the PBAC and were classified as comparator issues, comparative clinical efficacy issues, modeling issues, and calculation errors. This result was not surprising given that few governments were using HTA methods to inform their decision making. Therefore, the pharmaceutical industry had little need to generate the sort of clinical data needed by HTA agencies.

Like Hill et al, we found a large numbers of problems with the clinical evidence in submissions to the PBAC. The overall rate of significant problems in major submissions to the PBAC does not appear to have changed since the mid-1990s. We observed more problems with the analysis of the interpretation of the clinical evidence and fewer problems with the determination of therapeutic noninferiority than did Hill et al. This could be explained by a higher proportion of submissions for medicines with a claim of clinical superiority in our study. Because Hill et al. did not present any data to enable such a comparison, we are unable to determine this.

The findings from our study have international implications because it is likely that HTA agencies in other countries such as the National Institute for Health and Care Excellence have encountered similar types of problems in their assessment of the available clinical evidence for reimbursement/coverage determinations. A common denominator of the different publicly and privately funded health care technology reimbursement systems is the strength and relevance of comparative "clinical evidence." This holds true irrespective of whether or not the agencies require the evidence to be presented to them in submissions from the developers of the medicines concerned because the issue is more about the underlying clinical evidence rather than the failure of the developers to identify and then present it.

We are not aware of similar research being conducted in other jurisdictions, so we cannot determine whether our findings of the clinical evidence being a poor fit for purpose are confined to Australia. We note the recent studies by Kaltenthaler et al. [9,10] on the identification of issues associated with the first 95 single technology assessments undertaken by the National Institute for Health and Care Excellence and a review of the evidence to inform the population of cost-effectiveness models within HTAs. In neither study did they undertake an assessment of the quality of the supporting clinical evidence.

Problems with the choice of comparator are not unique to the PBAC; they have occurred with submissions to HTA agencies in other jurisdictions such as the Institute for Quality and Efficiency in Health Care in Germany and the Canadian Drug Expert Committee in Canada [11–13]. It is unclear as to whether choice of comparator issues are more or less common in Australia.

In the absence of empirical evidence on the quality of clinical evidence considered by other HTA agencies, the extent of problems in other jurisdictions is unknown.

The evaluation framework we developed could be used to conduct such research that could be performed by "independent" researchers if there is considerable information on the assessments and determinations in the public domain. Should such research be conducted and derive similar findings to ours, it will raise important issues regarding what can and should be done by all stakeholders to improve the quality of the clinical evidence used to support the reimbursement/coverage of new medicines.

Acknowledgments

We thank Sue Hill and Robyn Ward for their helpful comments on reviewing the manuscript.

Source of financial support: The authors have no financial relationships to disclose.

REFERENCES

[1] Drummond M, Neumann P, Jönsson B, et al. Can we reliably benchmark health technology assessment organizations? Int J Technol Assess Health Care 2012;28:159–65.
[2] Henry DA, Hill SR, Harris A. Drug prices and value for money: the Australian Pharmaceutical Benefits Scheme. JAMA 2005;294:2630–2.
[3] Hill SR, Mitchell AS, Henry DA. Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. JAMA 2000;283:2116–21.

[4] Australian Government, Department of Health and Ageing. Public summary documents explained. Available at: http://www.pbs.gov.au/info/industry/listing/elements/pbac-meetings/psd/pbac-psd-psd-explained. [Accessed February 20, 2015].
[5] Australian Government, Department of Health and Ageing, Immunise Australia Program. About the program. Available at: http://www.immunise.health.gov.au/internet/immunise/publishing.nsf/Content/about-the-program. [Accessed February 20, 2015].
[6] Australian Government, Department of Health and Ageing. Alternative arrangements for medicines. Other supply arrangements outside the Pharmaceutical Benefits Scheme: Life Saving Drugs Program. Available at: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pbs-general-supply-othersupply-copy2. [Accessed February 20, 2015].
[7] O'Brien B. Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? Med Care 1996;34(12 Suppl.):DS99–108.
[8] Rawlins M. Crossing the fourth hurdle. Br J Clin Pharmacol 2012;73:855–60.
[9] Kaltenthaler EC, Dickson R, Boland A, et al. A qualitative study of manufacturers' submissions to the UK NICE single technology appraisal process. BMJ Open 2012;2:e000562.
[10] Kaltenthaler E, Tappenden P, Paisley S. Reviewing the evidence to inform the population of cost-effectiveness models within health technology assessments. Value Health 2013;16:830–6.
[11] Ruof J, Schwartz FW, Schulenburg JM, Dintsios CM. Early benefit assessment (EBA) in Germany: analyzing decisions 18 months after introducing the new AMNOG legislation. Eur J Health Econ 2014;15:577–89.
[12] Rocchi A, Miller E, Hopkins RB, Goeree R. Common drug review recommendations: an evidence base for expectations? Pharmacoeconomics 2012;30:229–46.
[13] Rawlins M. *De testimonio*: on the evidence for decisions about the use of therapeutic interventions. Lancet 2008;372:2152–61.